View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by University of Queensland eSpace

INFECTION AND IMMUNITY, July 1995, p. 2409–2417 Vol. 63, No. 7
0019-9567/95/$04.00+0
Copyright © 1995, American Society for Microbiology

Clonal Relationships among Bloodstream Isolates of *Escherichia coli*

JOEL N. MASLOW,[1]* THOMAS S. WHITTAM,[2] CHARLES F. GILKS,[3] RICHARD A. WILSON,[3,4] MAURY ELLIS MULLIGAN,[5] KENNETH S. ADAMS,[1] AND ROBERT D. ARBEIT[1]

Infectious Diseases Section, Medical Service, VA Medical Center, and Department of Medicine, Boston University School of Medicine, Boston, Massachusetts[1]; Departments of Biology[2] and Veterinary Science,[4] The Pennsylvania State University, University Park, Pennsylvania; Infectious Diseases Section, Medical Service, VA Medical Center, Long Beach, California, and Department of Medicine, University of California, Irvine, California[5]; and Kenya Medical Research Institute, Nairobi, Kenya[3]

Received 23 February 1995/Returned for modification 17 March 1995/Accepted 5 April 1995

The clonal relationships among 187 bloodstream isolates of *Escherichia coli* from 179 patients at Boston, Mass., Long Beach, Calif., and Nairobi, Kenya, were determined by multilocus enzyme electrophoresis (MLEE), analysis of polymorphisms associated with the ribosomal operon (ribotyping), and serotyping. MLEE based on 20 enzymes resolved 101 electrophoretic types (ETs), forming five clusters; ribotyping resolved 56 distinct patterns concordant with the analysis by MLEE. The isolates at each study site formed a genetically diverse group and demonstrated similar clonal structures, with the same small subset of lineages accounting for the majority of isolates at each site. Moreover, two ribotypes accounted for ~30% of the isolates at each study site. One cluster contained the majority (65%) of isolates and, by direct comparison of the ETs and ribotypes of individual isolates, was genetically indistinguishable from the largest cluster for each of two other collections of *E. coli* causing pyelonephritis and neonatal meningitis (R. K. Selander, T. K. Korhonen, V. Väisänen-Rhen, P. H. Williams, P. E. Pattison, and D. A. Caugent, Infect. Immun. 52:213–222, 1986; M. Arthur, C. E. Johnson, R. H. Rubin, R. D. Arbeit, C. Campanelli, C. Kim, S. Steinbach, M. Agarwal, R. Wilkinson, and R. Goldstein, Infect. Immun. 57:303–313, 1989), thus defining a virulent set of lineages. The isolates within these virulent lineages typically carried DNA homologous to the adhesin operon *pap* or *sfa* and the hemolysin operon *hly* and expressed O1, O2, O4, O6, O18, O25, or O75 antigens. DNA homologous to *pap* was distributed among isolates of each major cluster, whereas *hly* was restricted to isolates of two clusters, typically detected in *pap*-positive strains, and *sfa* was restricted to isolates of one cluster, typically detected in *pap*- and *hly*-positive strains. The occurrence of *pap*-positive isolates in the same geographically and genetically divergent lineages suggests that this operon was acquired early in the radiation of *E. coli*, while *hly* and *sfa* were acquired subsequently, most likely by *pap*-positive and *pap*- and *hly*-positive precursors, respectively.

Escherichia coli is the most common gram-negative species causing bloodstream infections in humans. Most isolates of *E. coli* that infect the bloodstream possess virulence factors that allow the organisms to circumvent the normal clearance mechanisms, evade the host immune response, or have a selective growth advantage during pathogenesis (37, 43). Such virulence factors include adhesins which mediate attachment to mucosal epithelial surface glycoproteins and interfere with efficient surface clearance. Among these are P pili (encoded by the *pap* operon), S pili (encoded by the *sfa* operon), the afimbrial adhesin AFA (encoded by the *afa* operon), and the M afimbrial adhesin (encoded by the *bma* operon) (20, 24, 33, 47). Normal leukocyte function is compromised against bacteria that express hemolysin (encoded by *hly*) as well as against bacteria that express P pili (12, 54). In addition, macrophages demonstrate an altered cytokine response of interleukin 1α (IL-1α), IL-1β, IL-6, and tumor necrosis factor alpha when incubated with *E. coli* isolates that express P and S pili (1, 32). Bacteria that express the siderophore aerobactin, encoded by *aer*, are better able to replicate in environments with low iron content, such as the urine (41, 60). *E. coli* isolates from adult patients with bacteremia express these virulence factors and carry the respective virulence operons (37) at frequencies similar to those of *E. coli* isolates causing neonatal bacteremia, neonatal meningitis, bacteremia in children, and pyelonephritis in adults (2, 3, 25, 30, 56). In contrast, each of the operons encoding the respective virulence factors is detected less often among random fecal isolates (25, 34, 51).

The population structure of invasive extraintestinal isolates of *E. coli* has previously been examined in detail for pyelonephritic and neonatal meningitis isolates (4, 46, 50). Although individual collections of isolates recovered from invasive disease are genetically diverse, they represent a restricted subset of lineages within the overall population of *E. coli*, including random fecal isolates (8, 49, 51). Strains representing the most prevalent lineages typically express a combination of the virulence factors detailed above (4, 46, 50).

In this study, we determined the genetic diversity of 187 bloodstream isolates of *E. coli* cultured from patients at three diverse geographic sites (Long Beach, Calif.; Boston, Mass.; and Nairobi, Kenya). The genetic relatedness among isolates was inferred from protein polymorphisms as determined by multilocus enzyme electrophoresis (MLEE), examination of DNA polymorphisms detected by Southern blot analysis of the ribosomal operon, and analysis of antigenic variation revealed by serotyping. We also analyzed the genetic diversity of the adhesin operons *pap*, *sfa*, *afa*, and *bma*, the hemolysin operon *hly*, and the aerobactin operon *aer*.

MATERIALS AND METHODS

Patient populations and bacterial isolates. Bloodstream isolates of *E. coli* were obtained from patients at three different geographic sites. The collections represented all bloodstream isolates at each study site for the respective study

* Corresponding author. Mailing address: Research Service (151), VA Medical Center, 150 S. Huntington Ave., Boston, MA 02130. Phone: (617) 232-9500, ext. 5576. Fax: (617) 232-2969.

Downloaded from http://iai.asm.org/ on October 13, 2015 by University of Queensland Library

TABLE 1. Combinations of alleles for 20 enzyme loci that define 15 ETs represented by three or more isolates

Cluster	ET	Ribotype[a]	Study site[b]	No. of isolates	Allele at enzyme locus[c]													
					PGI	IDH	ACO	PE2	AK	MDH	PGD	M1P	GOT	BGA	ADH	MPI	TDH	SKD
I	62	Y	B, LB, N	4	5	2	7	5	2	4	5	6	6	8	4	4	2	3
III	5	E	B, LB	4	6	5	6	4	4	4	3	2	7	5	6	4	4	6
III	20	B2	B, LB	3	6	5	6	4	2	2	6	4	6	5	6	4	4	11
III	27	A	B	6	6	5	6	2.3	4	4	10	6	6	4.5	6	4	4	6
III	46	M	B, LB	3	6	5	6	4	2	4	4	6	7	6	6	6	4	6
III	68	B	B, LB	12	6	5	6	4	4	4	6	6	7	5	8	4	4	6
III	70	I, AT, BA	B, LB	6	6	5	6	4	4	4	6.8	2	7	5	6	4	4	6
III	71	D, AY	B, LB	5	6	5	5.5	4	4	4	6	2	6	5	6	6	4	6
III	72	A	B, LB	23	6	5	6	2.3	4	4	6	6	7	5	6	4	4	7
III	73	B2	B, LB, N	11	6	5	6	4	2	4	6	4	6	5	6	4	4	11
III	75	E	B, LB	4	6	5	6	4	4	4	10	2	7	4	6	4	4	6
III	79	E	B	3	6	5	6	4	4	4	4	2	7	4	6	4	4	8
III	96	A	N	3	6	5	6	2.3	4	4	6	6	7	5	6	4	4	6
IV	3	G	B, LB	3	6	5	6	5	2	4	6	8	6	4	6	10	4	8
V	2	AF	LB	5	7	5	5	5	2	6	6	7	6	4	6	6	4	11

[a] The ribotypes detected among the isolates correspond to those listed in Table 5 and presented in Fig. 1 and 2.
[b] Study sites are Boston (B), Long Beach (LB), and Nairobi (N).
[c] Only those 14 loci that demonstrate differences among the illustrated ETs are presented. The allelic migrations are the same as previously defined for other populations of *E. coli* (59). Abbreviations for enzyme loci: PGI, glucosephosphate isomerase; IDH, isocitrate dehydrogenase; ACO, aconitase; PE2, phenylalanyl-leucine peptidase; AK, adenylate kinase; MDH, malate dehydrogenase; PGD, gluconate-6-phosphate dehydrogenase; M1P, mannitol-1-phosphate dehydrogenase; GOT, aspartate aminotransferase; BGA, β-galactosidase; ADH, alcohol dehydrogenase; MPI, mannosephosphate isomerase; TDH, threonine dehydrogenase; SKD, shikimate dehydrogenase. Six loci were monomorphic among these 15 ETs: glyceraldehyde-3-phosphate dehydrogenase, glucose-6-phosphate dehydrogenase, indophenol oxidase, carbamate kinase, nucleoside phosphorylase, and glutamate dehydrogenase.

periods. The first collection comprised 119 isolates from 113 patients at the Boston VA Medical Center from January 1988 through December 1991. The second collection comprised 51 isolates from 50 patients at the Long Beach VA Medical Center from January through December 1991. The clinical and molecular epidemiologic characteristics of these isolates have been previously described in detail (37, 38). Four of these isolates were unavailable for the present analysis.

The third collection represented 21 isolates cultured from 20 patients at the Kenya Medical Research Laboratories/Wellcome Trust Research Laboratory, Nairobi, Kenya, from January 1989 through December 1991. Overall, a total of 187 isolates from 179 patients were included in this study. These 20 patients were identified as part of other studies, including a study of bacteremia among human immunodeficiency virus type 1 (HIV-1)-seropositive adults (three patients) (16), a prospective study of health care among prostitutes in Nairobi (eight patients, including Tanzanian women working in Nairobi), a study of spontaneous and induced abortions among women admitted to Kenyatta Hospital (three patients) (55), a study of the prevalence and morbidity due to infection with *Mycobacterium tuberculosis* (five patients) (9), and a study of acute pneumonia (one patient). Fourteen patients from Nairobi were female, 12 were HIV seropositive, and 5 had AIDS. The median age was 29.5 years (range, 19 to 38 years). In contrast, only two U.S. patients were HIV seropositive, and all were male; the median age was 65 years (range, 30 to 90 years) (37).

For each patient, the primary infecting site was determined through review of the clinical record. The primary sources of the bloodstream isolates were similar at each study site despite the notable differences in patient populations. Approximately 50% of isolates originated from a genitourinary focus of infection, 20% originated from a gastrointestinal focus, 5 to 10% originated from a pulmonary focus, and the remainder originated from other or unknown foci. Overall, 64% of the isolates represented community-acquired infections; 65% of the U.S. isolates and 100% of the African isolates represented community-acquired infections. The sites of infection also had similar frequencies: 63% of the genitourinary isolates and 60% of the gastrointestinal isolates represented community-acquired infections.

Isolates were collected to represent independent bacteremic episodes. Only one isolate per culture was analyzed except for two patients (one each at Boston and Nairobi) from whom two morphologically distinct isolates were identified. Both pairs of isolates represented distinct strains as determined by pulsed-field gel electrophoresis (data not shown) and were represented by different electrophoretic types (data not shown).

MLEE and phylogenetic analysis. We examined electrophoretic variation in 20 enzymes (Table 1) by the methods of MLEE as described previously (48, 59). Electromorphs of each enzyme, determined by comparison with standard mobility variants, were equated with alleles at the corresponding chromosomal gene locus. Distinctive multilocus genotypes (allele combinations) were designated electrophoretic types (ETs). Genetic distance (d) between ETs was calculated as $d_{ij} = -1n(1 - D)$, where D equals the proportion of loci with different alleles between the ith and jth ETs (21, 42). This assumes that any allelic difference in electrophoretic mobility reflects at least one codon difference at the nucleotide level and that codon changes occur independently; d then represents an estimate of the mean number of electrophoretically detectable codon differences per genetic locus. Genetic diversity (H), which is the probability that two isolates (ETs) differ at any particular locus, was calculated both within each study population and for the entire population (42). Phylogenetic relationships among ETs were inferred from cluster analysis by distance methods (15), using the average-linkage algorithm (52) to construct a dendrogram.

Isolation of whole cellular DNA and genotypic analyses. The isolation of chromosomal DNA in agarose plugs, preparation of DNA in solution, and subsequent restriction digestion have been described elsewhere (37, 39). All restriction enzymes were purchased from New England Biolabs (Beverly, Mass.) and used according to the manufacturer's directions. Restriction fragment length polymorphisms associated with the ribosomal operon (ribotypes) were determined by probing Southern blots of *Eco*RI-digested whole cellular DNA with the entire *E. coli rrnB* operon contained on a 7.5-kb *Bam*HI fragment of pC6 (6) as previously described (4, 36). On the basis of the patterns of restriction fragments, each strain was assigned a ribotype. Pattern names bear no relationship to each other (i.e., ribotypes A and AC and ribotypes B and B2 are independent types). Isolates assigned to the same ribotype were confirmed by comparison within the same gel.

The presence of each of the *E. coli* virulence operons *pap*, *sfa*, *afa*, *bma*, and *hly* was determined by probing dot blots of *Eco*RI-digested whole chromosomal DNA with intraoperonic fragments of each operon as previously described (37). The presence of *aer* was determined by probing with the 6.9-kb *Eco*RI-*Hin*dIII fragment of pABN1 spanning the entire aerobactin operon (pABN1 was kindly provided by Steven Opal) (7).

DNA hybridization. To prepare Southern blots, fragments of *Eco*RI-digested DNA were resolved in 0.8% agarose gels (Seakem GTG; FMC) by routine electrophoresis (Subcell; Bio-Rad) for 16 h at 35 V and transferred to a nylon membrane (Duralon; Stratagene, La Jolla, Calif.) under vacuum (Vacuum Blotter; Bio-Rad) at 5 mm Hg, baked at 80°C for 1 h, and then UV cross-linked to the membrane with 20,000 μJ (Stratalinker; Stratagene). Hybridization was performed overnight at 65°C in borosilicate tubes (hybridization chamber; Robbins Scientific, Sunnydale, Calif.), using 7% sodium dodecyl sulfate (SDS)–0.5 M NaH_2PO_4–1 mM EDTA, and then washed successively with 5% SDS–40 mM NaH_2PO_4–1 mM EDTA (two washes at 65°C for 30 min) followed by 1% SDS–40 mM NaH_2PO_4–1 mM EDTA (two washes at 65°C for 30 min). The DNA used as a probe was radiolabeled by using a random primer kit (Boehringer Mannheim, Indianapolis, Ind.) with [α-^{32}P]dCTP (NEN Research, Boston, Mass.) and separated from unincorporated nucleotides through a Sephadex G-50 Column (Nick Spin; Pharmacia, Uppsala, Sweden). The membranes were exposed to X-ray film (RX medical X-ray film; Fuji Photo Film, Elmsford, N.Y.) and processed in an automated film developer (RGII; Fuji Photo Film).

Serotyping. Analysis of the O and H serotypes was performed at the *E. coli* Reference Laboratory, Pennsylvania State University, University Park, as described previously (44). Agglutinations were performed with 173 standard World Health Organization antisera, 16 Ox (experimental) antisera, and 56 H antisera.

Downloaded from http://iai.asm.org/ on October 13, 2015 by University of Queensland Library

Strains that demonstrated agglutination with two or more O or H antisera were considered nontypeable.

RESULTS

MLEE of bloodstream isolates from three geographic sites. Analysis of the protein polymorphisms among the 187 isolates revealed extensive genetic variation in the genes encoding soluble enzymes. All 20 enzymes were polymorphic, with an average of 5.8 alleles per enzyme locus (range, 2 to 18). The average genetic diversity per locus within populations was 0.387, which is similar to values in other studies of diarrheagenic, invasive, and urinary tract isolates of *E. coli* (4, 50, 59). There were 101 distinct combinations of alleles (ETs); 15 ETs were represented by ≥3 isolates (range, 3 to 23 isolates) and accounted for 95 (51%) of all isolates (Table 1). Cluster analysis of the genetic distances among the 101 ETs resolved four clusters of lineages (I and III to V), with one isolate distantly related to these, designated cluster II (Fig. 1). Cluster III contained 122 (65%) of the study isolates and 12 of the 15 ETs represented by ≥3 isolates; the other three clusters contained ~20 isolates each (Table 2). The genetic diversities for the ETs within the various geographical populations were similar, with a slightly greater level of diversity in Nairobi (Table 3). Most loci were polymorphic, with the mean number of alleles per locus ranging from 3.4 for the Nairobi isolates to 5.1 for the Boston isolates (Table 3). Isolates at diverse geographic locales were represented by common lineages: of the 15 ETs with three or more isolates, 9 represented isolates from two study sites and 2 represented isolates at all three sites, accounting for 78 (42%) of study isolates (Table 1).

Evolutionary diversity associated with polymorphisms of the ribosomal operon. Among the 187 isolates, 56 distinct ribotypes were detected (Table 2). Nine ribotypes (16%) were represented by five or more isolates, 9 ribotypes (16%) were represented by two to four isolates, and 38 ribotypes (68%) were represented by a single isolate. Representative ribotypes are shown in Fig. 2. By mapping the analysis by ribotype onto that defined by ET, the genetic relatedness assessed by ribotyping and that assessed by MLEE were strongly correlated, as has previously been observed by others (4). Ribotypes represented by three or more isolates were often associated with several genetically related ETs (Fig. 1; Table 4); for those ribotypes comprising nine or more isolates, the average genetic diversity among ETs was 0.130 (Table 4). Only two ETs (70 and 71) included isolates represented by multiple ribotypes; both ETs contained individual isolates with unique ribotypes. Within each cluster, the majority of isolates were represented by a small number of ribotypes. For example, six ribotypes (A, B, B2, D, E, and H) accounted for 80% of the isolates in cluster III, two ribotypes (F and AF) accounted for 67% of isolates in cluster V, and one ribotype (G) accounted for 57% of isolates in cluster IV (Table 5). There was no correlation between the anatomic site of infection and either the ribotype or the ET (data not shown).

The geographic diversity of ribotypes was also consistent with the analysis by MLEE: 12 of the 14 ribotypes containing three or more isolates were detected at two or more study sites, and five ribotypes were detected at all three sites (Table 5). Two ribotypes, A and B2 (Fig. 2), accounted for ~30% of the isolates at each study sites (Table 5). Thus, a small number of lineages, as defined by MLEE and ribotyping, consistently accounted for the majority of study isolates and were typically detected at two or more study sites. A few ribotypes were found almost exclusively at a single site. Ribotype AF represented 12% of the isolates at Long Beach but none elsewhere; ribotype BE represented 15% of isolates at Nairobi but none elsewhere; and ribotype B represented 14% of the isolates at Boston but only one isolate at Long Beach.

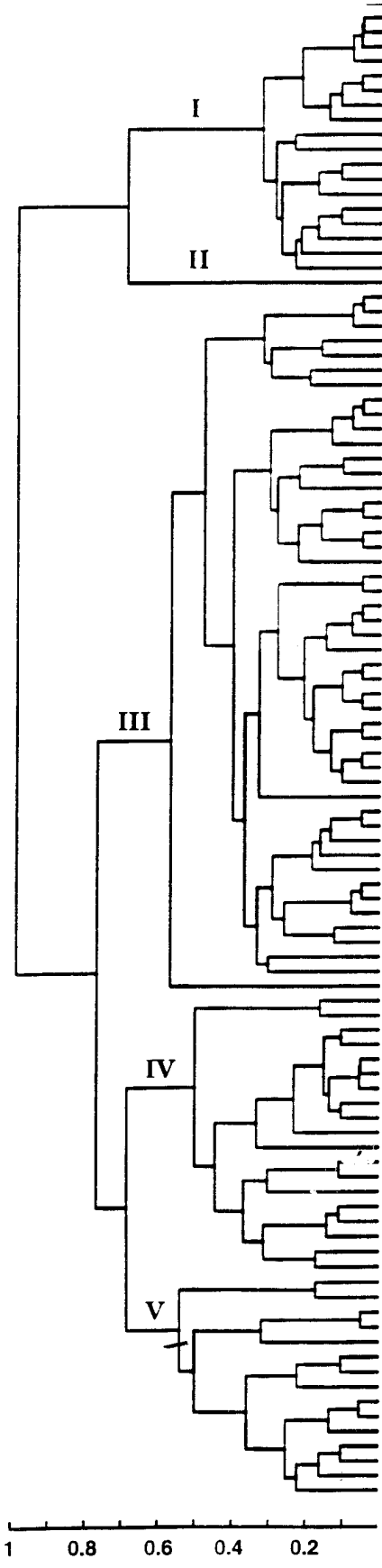
Correlation of serotype to lineage as defined by ET and ribotype. The O and H serotype data of the 187 isolates are presented in Fig. 1. The most common O groups were O1 (10 isolates), O2 (15 isolates), O4 (10 isolates), O6 (34 isolates), O8 (4 isolates), O15 (4 isolates), O18 (11 isolates), O25 (5 isolates), O75 (6 isolates), O85 (5 isolates), and O102 (6 isolates). Ten isolates agglutinated with two O antisera; 36 isolates were nontypeable or rough. There was poor typeability of these isolates with H antisera, as 51 isolates were nonmotile and 88 isolates were nontypeable with standard antisera. O:H serotypes were assigned to 43 isolates. Typically there was good correlation between O group, ribotype, and ET, with the majority of isolates of an ET expressing a common O antigen (Fig. 1). Also, while some O groups were restricted to a single MLEE cluster (e.g., isolates of serogroups O4 and O6 were found only in cluster III), other O groups were more genetically diverse (e.g., isolates of serogroups O2, O8, and O18).

Molecular epidemiology and geographic diversity of virulence factors. For each isolate, the presence of DNA homologous to the adhesin operons *pap*, *sfa*, *afa*, and *bma*, the α-hemolysin operon *hly*, and the aerobactin operon *aer* was assessed by probing with the respective intraoperonic probe. The frequency of each operon was similar for the African and U.S. isolates, none of the differences being statistically significant. *pap* was detected in 58 and 48% of the U.S. and Nairobi isolates, respectively. Similarly, *sfa* was detected in 37 and 24% of isolates, *afa* was detected in 6 and 24% of isolates, *bma* was detected in 2 and 0% of isolates, *hly* was detected in 44 and 29% of isolates, and *aer* was detected in 53 and 81% of isolates, respectively.

As previously described for the U.S. isolates (37), many of the African isolates were positive for multiple virulence factors. Overall, of the 107 *pap*-positive isolates, 79 (74%) were positive for sequences encoding another adhesin or hemolysin, including 54 *pap*, *sfa*, and *hly* positive and 19 *pap* and *hly* positive. Of the 67 *sfa*-positive isolates, 60 (90%) were positive for another adhesin or hemolysin. Of the 15 *afa*-positive isolates, 6 (40%) were positive for another adhesin or hemolysin. Each of the three *bma*-positive isolates was also *pap* positive. Each of the 79 isolates positive for *hly* was positive for one or more adhesins. Approximately 50% of the isolates positive for each adhesin or hemolysin were *aer* positive.

Genetic diversity of the virulence operons. There were significant differences in the frequencies with which virulence factors were present among isolates from different lineages (Tables 2 and 6). Although *pap*-positive isolates were detected in each of the four major clusters, 83% of all *pap*-positive isolates were included in cluster III and 11% of isolates were in cluster V. Within clusters III and V, 73 and 57% of isolates, respectively, were *pap* positive, compared with <15% of isolates within clusters I and IV (Table 2; $P < 0.0001$, χ^2 test). All 67 *sfa*-positive isolates were detected in cluster III ($P < 0.0001$, χ^2 test). *afa*-positive strains were detected in clusters I, III, and V; the highest frequency (27%) was for cluster I. *hly*-positive isolates were detected only within clusters III (60% of isolates) and V (29% of isolates; $P < 0.0001$, χ^2 test). Within each cluster, the isolates contained within the most highly represented lineages were predominately *pap*, *sfa*, or *hly* positive. For example, among the nine ribotypes represented by five or more isolates, 80% isolates were adhesin positive (Table 6), compared with 37% of isolates with unique ribotypes ($P < 0.0001$, χ^2 test). For three ribotypes (B, B2, and E), all of the isolates were adhesin positive.

Downloaded from http://iai.asm.org/ on October 13, 2015 by University of Queensland Library



Lineage	ET	Index strain	Ribotype(s)	Serotypes(s)	n	*pap*	*sfa*	*hly*	*afa*	*bma*	*aer*
						No. with virulence factor					
	62	Ca28	Y	U:U 12:U 101:9	4	2	-	-	1	1	3
	11	Ca36	AC	12:NM U:NM	2	-	-	-	2	-	2
	4	Ca21	AH	78:10	1	-	-	-	-	-	-
	88	Afr58	BE	U:NM	1	-	-	-	-	-	1
	89	Afr99	BE	U:NM	1	(1)	-	-	1	-	1
	35	Bos15	AA	40:U	1	-	-	-	-	-	1
	28	Bos109	AQ	11:52	1	-	-	-	1	-	1
	90	Afr36	BE	21:U	1	-	-	-	1	-	1
	86	Afr103	BL	U:U	1	-	-	-	-	-	-
	74	Ca6	BN	18:U	1	1	-	-	-	-	-
	56	Bos67	V	102:U	1	-	-	-	-	-	-
	32	Bos12	Z	15:U	1	-	-	-	-	-	-
	103	Ca46	AJ	U:U	1	-	-	-	-	-	-
	60	Bos81	X	8:U	1	-	-	-	-	-	-
	33	Bos121	AS	U:NM	1	-	-	-	-	-	-
	1	Ca1	BQ	7:18	1	-	-	-	-	-	1
	87	Afr96	BH	U:NM	1	-	-	-	-	-	-
	57	Bos76	W	75:U	1	-	-	-	-	-	-
	102	Afr102	BK	U:NM	1	-	-	-	-	-	-
	17	Ca55	H	107:U	1	1	-	-	-	-	1
	6	Ca25	H	U:56	1	1	-	(1)	-	-	1
	16	Bos77	H	25:U	2	1	-	-	-	-	2
	61	Bos82	H	85/173:U	1	-	-	-	-	-	-
	59	Bos79	H	U:U	2	-	-	-	-	-	-
	84	Bos56	AZ	25:U	1	-	-	-	-	-	-
	42	Bos32	BP	85:U	1	1	-	-	-	-	1
	67	Bos95	C	16:U	1	1	-	1	-	-	1
	30	Bos114	K	16:U	1	1	-	1	-	-	-
	82	Bos10	C	16:48	1	1	-	1	-	-	1
	99	Afr26	BS	U:U	1	1	-	-	-	-	1
	36	Bos23	N	85/157:5	1	-	1	-	-	-	-
	13	Ca44	B2	U:NM 21:NM	2	2	1	2	1	-	2
	66	Bos92	BU	81:NM	1	-	-	-	-	-	-
	73	Ca2	B2	4:U 4:NM 4:5 4/18:U 21:U 85:U	10	9	7	10	2	-	5
	20	Bos38	B2	4:U 4:NM 4:5	3	3	3	3	-	-	1
	65	Bos88	BS	U:NM	1	-	1	1	-	-	-
	55	Bos64	BS	2:4	1	-	1	-	-	-	-
	38	Bos25	BR	21/83:U	1	1	1	1	-	1	-
	27	Bos89	A	2:U 2:NM	6	6	4	5	-	-	6
	9	Ca32	A	2:U	1	1	1	1	-	-	1
	69	Bos65	B	6:NM	1	1	-	1	-	-	-
	68	Bos72	B	6:U 6:NM 6:16 6:29 6:38/56	12	12	11	12	-	-	-
	81	Bos69	B	6:U	1	1	1	1	-	-	-
	29	Bos111	B	85:U	1	1	1	1	-	-	-
	43	Bos33	A	U:NM	1	-	1	-	-	-	1
	21	Ca7	A	18:U	1	-	-	-	-	-	1
	72	Bos19	A	6:U 6:1 6:16, 6:16/18 25:U	23	14	15	17	-	-	19
	58	Bos78	A	U:4/32	1	-	1	-	-	-	-
	98	Afr104	A	U:U	1	1	1	1	-	-	1
	96	Afr100	A	6:44 6:32/44 U:16/18	3	3	3	1	-	-	3
	70	Ca14	I, BA, AT	18:U 18:NM 18:–7 18:15	6	6	6	6	-	-	2
	5	Bos29	E	1:–, 1:24, 1:NM	4	4	-	-	-	-	3
	31	Bos117	B	6:16/18	1	1	1	1	-	-	-
	85	Bos62	T	85:1	1	-	-	-	-	-	-
	51	Bos48	D	18:NM	2	2	1	2	-	-	2
	48	Bos44	D	18:NM	1	1	1	1	-	-	1
	100	Afr105	D	75:U	1	1	-	-	-	-	1
	71	Bos66	D, AY	75:U 75:NM 75/138:NM	5	1	1	1	3	-	4
	83	Bos51	D	18:NM	1	1	1	1	-	-	1
	80	Bos112	E	2:U	1	1	-	-	-	-	-
	75	Ca57	E	2:U 2:NM	4	4	-	-	-	-	-
	79	Bos6	E	1:U 1:7	3	3	-	-	-	-	3
	50	Bos5	B	6:U	1	1	1	1	-	-	-
	44	Bos37	D	13w:U	1	-	1	-	-	-	-
	46	Ca16	M	46/171:U 171:U	3	-	-	-	-	-	-
	8	Ca29	BT	120:14	1	-	-	-	-	-	-
	37	Bos24	BM	U:U	1	-	-	-	-	-	-
	101	Afr70	BF	23:NM	1	-	-	-	-	-	-
	64	Bos87	G	86:U	1	-	-	-	-	-	1
	63	Bos85	G	58:40	1	-	-	-	-	-	1
	41	Bos3	G	8:U	1	-	-	-	-	-	1
	34	Bos14	G	158:U	1	1	-	-	-	-	1
	25	Bos103	G	U:U	1	-	-	-	-	-	-
	53	Bos59M	G	U:NM	1	-	-	-	-	-	-
	22	Ca9	G	U:U	1	-	-	-	-	-	-
	3	Bos99	G	8:12 U:U U:NM	3	-	-	-	-	-	1
	26	Bos106	G	85:U	1	-	-	-	-	-	-
	23	Bos1	G	8:19	1	-	-	-	-	-	-
	52	Bos59	Q	2:4	1	-	-	-	-	-	1
	49	Bos47	Q	U:U	1	-	-	-	-	-	-
	95	Afr22	BD	U:U	1	-	-	-	-	-	1
	76	Bos36	J	U:NM	1	1	-	-	-	-	1
	10	Ca33	AE	U:U	1	1	-	-	-	-	1
	40	Bos28	P	20:U	1	-	-	-	-	-	1
	15	Ca52	AL	166:U	1	-	-	-	-	-	1
	12	Ca37	AN	11:52	1	-	-	-	-	-	1
	77	Ca40	AF	2:U	1	1	-	1	-	-	1
	2	Ca50	AF	102:56	5	2(2)	-	2	1	-	4
	39	Bos 26	O	86:U	1	-	-	-	-	-	-
	7	Ca26	O	68/73:U	1	-	-	-	-	-	1
	54	Bos63	U	51:U	1	-	-	-	-	-	-
	18	Ca56	F	82/85:NM	1	1	-	-	-	-	1
	14	Ca49	F	7:NM	1	1	-	(1)	-	1	1
	93	Afr97	BI	U:NM	1	-	-	-	1	-	1
	91	Ca17	F	15:U	1	1	-	1	-	-	1
	92	Afr17	F	15:U	2	2	-	2	-	-	2
	19	Ca59	AW	15:7	1	1	-	-	-	-	1
	78	Bos58	F	1:NM	2	1(1)	-	-	-	-	2
	45	Bos39	F	1:NM	1	1	-	-	-	-	1
	47	Bos43	AD	157:U	1	1	-	-	1	-	1
	24	Bos102	AV	86:U	1	-	-	-	-	-	-

Downloaded from http://iai.asm.org/ on October 13, 2015 by University of Queensland Library

FIG. 1. Genetic relationships of 187 *E. coli* bloodstream isolates from Boston, Long Beach, and Nairobi as resolved by MLEE. Cluster analysis revealed 101 distinct ETs. Cluster I includes lineages 1 to 18, cluster II includes lineage 19 only, cluster III includes lineages 20 to 67, cluster IV includes lineages 68 to 86, and cluster V includes lineages 87 to 101. One index isolate is listed for each ET. The genetic distance, in increments of 0.1, is indicated below the dendrogram. The ribotypes represented by the isolates within each ET are given; only the most prevalent serotypes are listed. Abbreviations for serotype designations: U, untypeable (this category includes both nonagglutinating and multiply agglutinating isolates); NM, nonmotile; w, weak reaction. Prefixes for sites: Ca, Long Beach; Afr, Nairobi; Bos, Boston. The number of isolates with partial copies of *pap* or *hly* are included in parentheses. *papEFG* only was detected for three isolates in ETs 2 and 89, *papHCD* only was detected for one isolate in ET 78, and *hlyA* only was detected for two isolates in ETs 6 and 14.

DISCUSSION

E. coli is the most common gram-negative species that causes septicemia and extraintestinal infections such as meningitis and pyelonephritis. Prior studies using MLEE to examine the clonal structures of the *E. coli* isolates that cause neonatal meningitis and pyelonephritis have observed that such isolates represent a relatively limited set of genetic lineages compared with the population of fecal flora (4, 11, 49, 50). These studies confirmed the clonal hypothesis suggested by examination of the O:K:H serotypes of invasive strains (57). DNA-based analyses indicated that these isolates typically contain sequences encoding for adhesins (e.g., *pap*, *sfa*, and *afa*), hemolysin (*hly*), or aerobactin (3, 5, 30, 37) and express virulence factors involved in pathogenesis. In contrast, such factors are found in only 10 to 15% of fecal isolates from normal hosts (10, 26).

In this study, we assessed the genetic diversity and clonal relationships of *E. coli* isolates infecting the bloodstream and extended our previous study of clinical isolates from Long Beach, Calif., and Boston, Mass. (37), to include organisms from Nairobi, Kenya. The analysis yielded a number of key findings. (i) The isolates from each of the three geographically diverse locations were genetically diverse, but the clonal structures of the isolates were similar. (ii) Among the genetic lineages resolved by MLEE, ribotyping, and serotyping, the majority of isolates at each study site represented a small subset of lineages. (iii) The frequently represented lineages were the same at each locality, thus defining a set of widespread invasive clones. Most of these lineages were genetically related, i.e., were contained within the same cluster. (iii) The isolates represented by these invasive clones typically contained one or more adhesins or hemolysin.

Clonal structures of *E. coli* bloodstream isolates. Examination of the genetic relationships among the 187 bloodstream isolates by MLEE of 20 enzyme loci, ribotyping, and serotyping resolved a large number of lineages. By mapping the analysis by ribotype onto the analysis by ET, we found that MLEE was more discriminatory than ribotyping, with multiple ETs often detected within individual ribotypes, similar to the observation of other investigators (4). Both analyses defined similar genetic relationships; i.e., the genetic distance among ETs representing a single ribotype was typically ≤ 0.2. The clonal structures of isolates were similar at each study site; that is, the isolates were genetically diverse, but few lineages accounted for the majority of isolates. Of note, the most common lineages were the same for each study site, defining a set of virulent clones. A single ribotype (ribotype A) comprised a genetically homogeneous ($H = 0.082$) group of eight ETs and accounted for ~20% of the isolates at each study site; a second ribotype (B2) accounted for ~10% of the isolates at each site. The majority of virulent lineages that were genetically related and a single cluster (cluster III) contained the majority (65%) of study isolates. Further, 86 of the cluster III isolates expressed O antigens typically found among isolates causing pyelonephritis (e.g., O1, O2, O4, O6, O7, O8, O16, O25, and O75) (28, 56); these isolates represented 85% of all isolates expressing these O antigens. None of the six isolates expressing O7 and O8 antigens were included in cluster III.

Other studies using MLEE or serotyping have found that the majority of *E. coli* isolates from urinary tract infections or neonatal meningitis similarly represent a relatively few, genetically related lineages (4, 46, 50, 57). Plos et al. have further noted that isolates from Goteberg, Sweden, and Houston, Tex., may be contained within the same lineages (46).

We directly compared the electrophoretic profiles of 20 metabolic enzymes for selected isolates from this study with those of isolates from several previously defined collections. Selander et al. examined *E. coli* from infants in Finland with meningitis or septicemia (50); isolates representing ETs 9 and 18 of that study were indistinguishable from, and therefore clonally related to, ETs 72 and 73, respectively, as reported here. In addition, since isolates corresponding to subgroups A, B, and C of that study were present in our collection from North America and Africa, we conclude that our cluster III directly corresponds to group 1 of Selander et al., which represented 63% of the Finnish isolates. Arthur et al. examined bloodstream isolates from 23 patients from Boston with urinary and biliary tract sepsis (4); isolates from 5 of those patients were included among the sequential bloodstream isolates reported here (isolates 12 [cluster I, ET 32], 23 [cluster III, ET 36], 25 [cluster III, ET 38], 26 [cluster V, ET 39], and 27 [cluster III, ET 73], respectively). Thus, group IV of Arthur

TABLE 2. Correlation of electrophoretic clusters with ribotype and prevalence of virulence factors

Cluster	No. (%) of isolates	No. of ETs	No. of ribotypes	% of isolates within each cluster with virulence factor						
				Any adhesin	*pap*	*sfa*	*afa*	*bma*	*hly*	*aer*
I	22 (12)	18	16	41	14[a]	0	27	5	0	50
II	1 (1)	1	1	0	0	0	0	0	0	0
III	122 (65)	48	22	85	73	55	5	1	60[b]	53
IV	21 (11)	19	9	14	14	0	0	0	0	57
V	21 (11)	15	8	62	57[a]	0	14	5	29[b]	81
Total	187 (100)	101	56	69	57	36	8	2	42	56

[a] Partial copies of the *pap* operon were detected for four additional isolates. One isolate in cluster V was positive for *papHCD* only; one isolate in cluster I and two isolates in cluster V were positive for *papEFG* only.

[b] Partial copies of the *hly* operon were detected for additional two isolates; one isolate each in clusters III and V was positive for *hlyA* only.

Downloaded from http://iai.asm.org/ on October 13, 2015 by University of Queensland Library

TABLE 3. Genetic diversity among *E. coli* bloodstream isolates from three geographically diverse study sites

Study site	No. of:		% Polymorphic loci	Mean no. of alleles/locus	ETs	Isolates
	Isolates	ETs			$H \pm$ SE	$H \pm$ SE
Boston	115	67	100	5.1	0.402 ± 0.064	0.340 ± 0.061
Long Beach	51	33	85	4.0	0.418 ± 0.069	0.386 ± 0.067
Nairobi	21	16	85	3.4	0.460 ± 0.062	0.436 ± 0.060

et al. also corresponds to our cluster III. Taken together, these results indicate that isolates representing a single cluster of genetic lineages have been identified by multiple studies as a major cause of extraintestinal infections on three continents.

Some lineages representing multiple isolates were detected at only a single study site (e.g., ribotype BE was detected only at Nairobi, while ribotype AF was detected only at Long Beach). These episodes of bacteremia are unlikely to represent infection from a common source since, except for the five patients with episodes of recurrent bacteremia (38), isolates within the same lineage represented distinct strains as resolved by pulsed-field gel electrophoresis (unpublished results). Thus, some lineages may have undergone more recent divergence.

Clonal structures of virulence factor operons among *E. coli* bloodstream isolates. *pap* and *sfa* were the two most common virulence factor operons carried by the U.S. (37) and African isolates. *pap* was detected in ~55% of isolates, *sfa* was detected in ~30% of isolates, *afa* was detected in ~10% of isolates, *hly* was detected in ~40% of isolates, and *aer* was detected in ~55% of isolates. Despite appreciable differences in patient populations at the three study sites, there were only minor differences in the frequencies of virulence factors.

As noted above, cluster III contained the largest fraction of study isolates; the majority of these isolates carried the virulence operons *pap*, *sfa*, and *hly*. Of the 122 isolates within cluster III, 101 (83%) contained DNA homologous to at least one of these operons: 73% were *pap* positive, 60% were *hly* positive, and 55% were *sfa* positive. These virulence factor operons were much less frequent among the isolates representing the other clusters. In contrast, *aer* was present in ~55% of the isolates in each cluster.

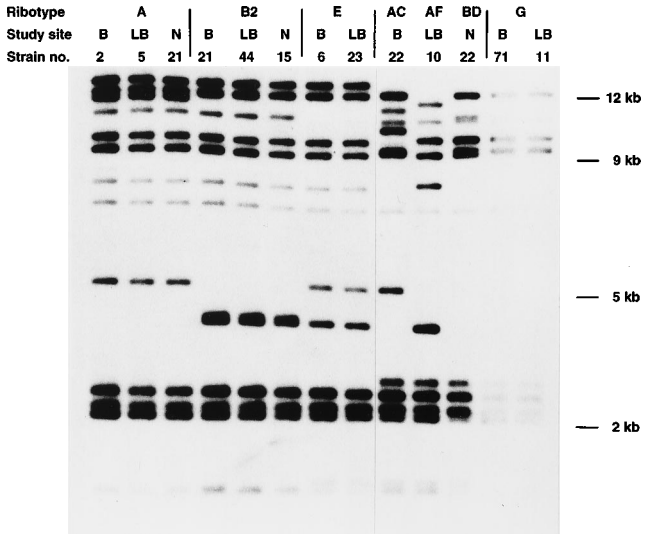


FIG. 2. Representative ribosomal polymorphisms (ribotypes) for isolates from the three study sites. Ribotype A was the most prevalent pattern detected among the isolates at each of the three sites; ribotype B2 was the second most prevalent pattern. Ribotype G was detected in 7% of all study isolates. Abbreviations: B, Boston; LB, Long Beach; N, Nairobi.

TABLE 4. Genetic diversity within ribotypes

Ribotype	No. of isolates	No. of ETs	$H \pm$ SE
A	37	8	0.082 ± 0.033
B	17	6	0.053 ± 0.016
B2	15	3	0.030 ± 0.019
D	10	6	0.151 ± 0.054
E	13	5	0.102 ± 0.048
G	12	10	0.185 ± 0.057
F	7	6	0.238 ± 0.062
H	7	5	0.200 ± 0.066
AF	6	2	0.067 ± 0.031

In addition to these well-known virulence factors, certain genotypes may have other characteristics that facilitate the development of invasive infection. This hypothesis is supported by two findings, notwithstanding the observation that most patients in this study infected with adhesin-negative isolates had identified defects in normal mucosal barriers or defects in immunity (37): (i) one lineage (ribotype G) found among multiple patients at different study sites contained predominately (93%) adhesin-negative isolates, and (ii) some lineages (e.g., ribotype A) represented both virulence factor-positive and virulence factor-negative isolates. In addition, only one of the five patients in this study with recurrent bacteremia was infected with adhesin-negative isolates; these isolates represented ribotype G (38). We have also identified adhesin operon- and *hly*-negative uropathogenic isolates representing ribotype G cultured from patients in Des Moines, Iowa; Memphis, Tenn.; Denmark; and Sweden (unpublished results). Ribotype G isolates do not share a common serogroup, nor do they frequently carry other virulence factors (e.g., aerobactin). To our knowledge, this ribotype represents the first predominately adhesin operon-negative lineage that is widely represented among collections of invasive isolates from multiple geographic sites.

Although the ribotype G isolates did not express mannose-resistant hemagglutination to either sheep or human erythrocytes (37), other adhesins that bind to either carbohydrate or protein moieties of other cell types may be expressed on the bacterial surface and give these isolates a selective colonization advantage. In addition, the possibility exists that these isolates express another cytolysin that allows for tissue invasion or survival within leukocytes. Thus, this group of organisms may express factors, as yet unidentified, that provide for tissue colonization, translocation, or serum resistance.

Evolutionary and genetic basis of the distribution of virulence operons. The detection of *pap* among the same genetically diverse lineages worldwide suggests that this operon was a phylogenetically distant acquisition by *E. coli* with little horizontal transfer. Although horizontal movement of *pap* has been postulated to occur for this operon (4, 46), there is still, to our knowledge, no molecular data to support this view directly. Our extensive data set suggests a possible alternative explanation, specifically, that after acquisition, *pap* was spontaneously lost from individual ancestral *E. coli* clones. The lineages that retained *pap* would then have a selective advantage in causing invasive disease and thus be detected in septic patients more frequently. Spontaneous deletion of *pap* and *hly* has been documented in vitro at low frequencies for some clinical isolates (17, 18). It is possible that those *pap*-positive

Downloaded from http://iai.asm.org/ on October 13, 2015 by University of Queensland Library

TABLE 5. Geographic diversity of the 14 ribotypes representing three or more isolates

Cluster	Ribotype[a] (*n*)	ETs represented	% of isolates at study site represented by ribotype			
			Boston (*n* = 115)	Long Beach (*n* = 51)	Nairobi (*n* = 21)	Total (*n* = 187)
I	Y (4)	62	1	2	10	2
I	BE (3)	88, 89, 90	0	0	15	2
III	A (37)	9, 21, 27, 43, 58, 72, 96, 98	20	20	19	20
III	B (17)	29, 31, 50, 68, 69, 81	14	2	0	9
III	B2 (15)	13, 30, 73	8	8	10	8
III	D (10)	44, 48, 51, 71, 83, 100	5	6	5	5
III	E (12)	5, 75, 79, 80	7	8	0	6
III	H (7)	6, 16, 17, 59, 61	4	4	0	4
III	I (4)	70	2	4	0	2
III	M (3)	46	2	2	0	2
III	BS (3)	55, 65, 99	2	0	5	1
IV	G (12)	3, 22, 23, 25, 26, 34, 41, 53, 63, 64	9	4	0	6
V	F (8)	14, 18, 45, 78, 91, 92	3	6	10	4
V	AF (6)	2, 77	0	12	0	3

[a] Each distinct ribotype pattern was assigned a unique arbitrary designation; isolates with patterns that differed by one or two bands consistent with a single genetic event are considered to represent a common ribotype.

lineages detected at a single study site (e.g., ribotype AF) represent recent horizontal transfer of this operon.

The *hly* operon was detected among only isolates representing two clusters (clusters III and V) and only in adhesin-positive strains (93% were *pap* positive, and 73% were *sfa* positive). The association of *hly* with *pap* and *sfa* has been well documented (3, 50). In addition, *pap* and *hly* are closely linked on the *E. coli* chromosome in all isolates studied (17, 22, 23, 35, 45). These data suggest that *hly* was acquired into a *pap*-positive background. The hypothesis that the *hly* operon was imported into *E. coli* from another species is suggested by the finding that the codon usage of *hly* is atypical for *E. coli* genes. Given the low GC content of this operon (39%) (14) and the high degree of sequence homology to the *Proteus vulgaris* RTX hemolysin (31) and the *Pasteurella haemolytica* leukotoxin (53), both species with a GC content similar to that of *hly* (39 and 43%, respectively), either a *Proteus* or a *Pasteurella* species may have been the source of this virulence factor. In some isolates, *hly* is flanked by insertion elements (40), suggesting a possible mechanism for such transfers. If horizontal transfer of *hly* within *E. coli* has occurred, then it is likely to have moved in concert with *pap*. The observation of different *pap-hly* linkage distances (35) may represent either multiple independent transfers of *hly* into a *pap*-positive progenitor or recombination within a single strain. Independent horizontal transfer events would be supported by consistent linkages for *pap* and *hly* only among isolates within individual clusters.

TABLE 6. Prevalence of DNA homologous to the operons encoding virulence factors among the ribotypes represented by five or more isolates

Ribotype (*n*)	% of isolates positive within each ribotype positive for virulence factor						
	Any adhesin	*pap*	*sfa*	*afa*	*bma*	*hly*	*aer*
A (37)	86	68	70	0	0	66	86
B (17)	100	100	88	0	0	100	0
B2 (15)	100	93	73	20	0	100	53
D (10)	90	60	50	20	0	50	80
E (12)	100	100	0	0	0	0	50
F (8)	88	88[a]	0	0	13	38[b]	100
G (12)	8	8	0	0	0	0	42
H (7)	43	43	0	0	0	0[b]	57
AF (6)	50	50[a]	0	17	0	50	83

[a] Partial copies of the *pap* operon were detected for four additional isolates. One isolate represented by ribotype F was positive for *papHCD* only; two isolates represented by ribotype AF and one isolate represented by ribotype BE were positive for *papEFG* only.

[b] Partial copies of the *hly* operon were detected for two additional isolates; one isolate each represented by ribotypes F and H was positive for *hlyA* only.

The *sfa* operon was detected only among isolates of cluster III and usually in association with *pap* and/or *hly* (90% of *sfa*-positive isolates were *pap* positive, and 86% were *hly* positive). The association between these operons is similar to that described by other investigators (3, 50). For *E. coli* isolate 536, *sfa* has been reported to map to the same general region of the chromosome as *pap* and *hly* (29); nevertheless, *sfa* was present on a distinct ~400-kb *Xba*I fragment (17, 45), separated by ≥190 kb from the other virulence factors (19, 29). Of note is that strain 536 has a single allele difference from our isolate Long Beach 24 (ET 72, cluster III), indicating that this strain belongs to the cluster III group (58a). Overall, the available data suggest that *sfa* was acquired relatively recently by a *pap*- and *hly*-positive progenitor and has undergone little, if any, subsequent horizontal movement.

The small numbers of *afa*- and *bma*-positive isolates preclude an in-depth analysis of the mechanisms of genetic diversity. *afa* was detected among 15 isolates from three clusters (I, III, and V), with no consistent relationship to any other virulence factor. *afa* also appears to have been phylogenetically distant acquisition by *E. coli*, as suggested by a number of observations. First, the *afa*-positive isolates from different study sites (Long Beach and Boston) were represented by the same ETs (ETs 11 and 71). Second, the three *afa*-positive isolates reported by Arthur et al. (4) were contained in lineages analogous to those defined in the present study. Mapping studies using pulsed-field gel electrophoresis show no evidence of physical linkage of *afa* or *bma* to *pap* or *hly* on the *E. coli* chromosome (unpublished data).

In contrast to the chromosome-associated virulence factor operons discussed above, *aer* can be commonly found on either colicin V, IncFI plasmids, or the chromosomes of isolates of *E. coli* cultured from humans (13, 60). In our series, the distribution of *aer* was nonrandom and appeared to be clonal. For example, some ribotypes contained ≥80% *aer*-positive isolates (ribotypes A, D, F, and AF), whereas others (ribotype B) had

Downloaded from http://iai.asm.org/ on October 13, 2015 by University of Queensland Library

no *aer*-positive isolates. Such clonal distribution of *aer* has also been reported by others (58). Plasmid versus chromosomal carriage of *aer* has been roughly correlated with the O group and membrane protein pattern (58) and has been reported to be inversely related to chromosomal carriage of *pap* (27). An analysis of whether *aer* was episomal or chromosomally integrated among our isolates would further define the clonal structure of this virulence factor.

In conclusion, despite the considerable genetic diversity among all isolates of *E. coli* causing bloodstream infection, pyelonephritis, and neonatal septicemia and pyelonephritis, a subset of lineages accounts for the majority of isolates, and the same virulent clones are detected at geographically diverse sites. These pathogenic lineages have an appreciably increased frequency of the virulence operons *pap*, *sfa*, and *hly*, consistent with the hypothesis that the encoded virulence factors contribute to invasiveness. Acquisition of these virulence operons appears to be sequential, with *pap* acquired prior to *hly* and both virulence operons acquired prior to *sfa*. Determination of the relative contribution of horizontal transfer to the genetic diversity of the individual operons will require more detailed studies, including chromosomal mapping and sequence analysis of virulence factors among individual strains representing diverse genetic lineages.

ACKNOWLEDGMENTS

We are grateful to Werner Goebel, Sheila Hull, Staffan Normark, Steven Opal, Mikael Rhen, Rodney Welch, David Hasty, and Evgeni Sokurenko for gifts of bacterial strains and plasmids; Stephen Brecher, Annette Durbin, Diane Fitzsimmons, Elisa Tosi, Susan Loring, Sue Ellen Briggs at the Boston VAMC and D. Bozanic, M. Collins, G. Holmes, D. Pham, and R. Voll from the Long Beach VAMC for the collection of bacterial strains; Janice Justis for clinical data collection; Michael Davis, Brian Miller, and Marie Wolfe at Pennsylvania State University for technical assistance; Mohamed Amir, Richard Brindle, Barry Batchelor, Joseph Kimari, Sam Kariuki, Marleen Temmerman, and staff and colleagues in the Kenya Medical Research Institute, Nairobi, for technical assistance; and Marie Scurti for assistance in the preparation of the manuscript.

Financial support was received from the Medical Research Program, Department of Veterans Affairs, Washington, D.C. (J.N.M. and R.D.A.); the Wellcome Trust (C.F.G.); and Public Health Service grants AI 24566 and RCDA AI 00964 (T.S.W.).

REFERENCES

1. **Agace, W., S. Hedges, U. Andersson, J. Andersson, M. Ceska, and C. Svanborg.** 1993. Selective cytokine production by epithelial cells following exposure to *Escherichia coli*. Infect. Immun. **61:**602–609.
2. **Archambaud, M., P. Courcoux, and A. Labigne-Roussel.** 1988. Detection by molecular hybridization of pap, afa, and sfa adherence systems in *Escherichia coli* strains associated with urinary and enteral infections. Ann. Inst. Pasteur Microbiol. **139:**575–588.
3. **Archambaud, M., P. Courcoux, V. Ouin, G. Chabanon, and A. Labigne-Roussel.** 1988. Phenotypic and genotypic assays for the detection and identification of adhesins from pyelonephritic *Escherichia coli*. Ann. Inst. Pasteur Microbiol. **139:**557–573.
4. **Arthur, M., R. D. Arbeit, C. Kim, P. Beltran, H. Crowe, S. Steinbach, C. Campanelli, R. A. Wilson, R. K. Selander, and R. Goldstein.** 1990. Restriction fragment length polymorphisms among uropathogenic *Escherichia coli* isolates: *pap*-related sequences compared with *rrn* operons. Infect. Immun. **58:**471–479.
5. **Arthur, M., C. E. Johnson, R. H. Rubin, R. D. Arbeit, C. Campanelli, C. Kim, S. Steinbach, M. Agarwal, R. Wilkinson, and R. Goldstein.** 1989. Molecular epidemiology of adhesin and hemolysin virulence factors among uropathogenic *Escherichia coli*. Infect. Immun. **57:**303–313.
6. **Berg, K. L., C. L. Squires, and C. Squires.** 1987. In vivo translation of a region within the *rrnB* 16S rRNA gene of *Escherichia coli*. J. Bacteriol. **169:**1691–1701.
7. **Bindereif, A., and J. B. Neilands.** 1983. Cloning of the aerobactin-mediated iron assimilation system of plasmid ColV. J. Bacteriol. **153:**1111–1113.
8. **Brauner, A., J.-M. Boeufgras, S. H. Jacobson, B. Kaijser, G. Källenius, S. B. Svenson, and B. Wretlind.** 1987. The use of biochemical markers, serotype and fimbriation in the detection of *Escherichia coli* clones. J. Gen. Microbiol. **133:**2825–2834.
9. **Brindle, R. J., P. P. Nunn, B. I. F. Batchelor, S. N. Gathua, J. N. Kimari, R. S. Newnham, and P. G. Waiyaki.** 1993. Infection and morbidity in patients with tuberculosis in Nairobi, Kenya. AIDS **7:**1469–1474.
10. **Carbonetti, N. H., S. Boonchai, S. H. Parry, V. Väisänen-Rhen, T. K. Korhonen, and P. H. Williams.** 1986. Aerobactin-mediated iron uptake by *Escherichia coli* isolates from human extraintestinal infections. Infect. Immun. **51:**966–968.
11. **Caugant, D. A., B. R. Levin, G. Lidin-Janson, T. S. Whittam, C. Svanborg-Edén, and R. K. Selander.** 1983. Genetic diversity and relationships among strains of *Escherichia coli* in the intestine and those causing urinary tract infections. Prog. Allergy **33:**203–227.
12. **Cavalieri, S. J., and I. S. Snyder.** 1982. Effect of *Escherichia coli* alpha-hemolysin on peripheral leukocyte function in vitro. Infect. Immun. **37:**966–974.
13. **Colonna, B., L. Ranucci, P. A. Fradiani, M. Casalino, A. Calconi, and M. Nicoletti.** 1992. Organization of aerobactin, hemolysin, and antibacterial resistance genes in lactose-negative *Escherichia coli* strains of serotype O4 isolated from children with diarrhea. Infect. Immun. **60:**5224–5231.
14. **Felmlee, T., S. Pellett, and R. A. Welch.** 1985. Nucleotide sequence of an *Escherichia coli* chromosomal hemolysin. J. Bacteriol. **163:**94–105.
15. **Felsenstein, J.** 1984. Distance methods for inferring phylogenies: a justification. Evolution **38:**16–24.
16. **Gilks, C. F., R. J. Brindle, L. S. Otieno, P. M. Simani, R. S. Newnham, S. M. Bhatt, G. N. Lule, G. B. A. Okelo, W. M. Watkins, P. G. Waiyaki, J. B. O. Were, and D. A. Warrell.** 1990. Life-threatening bacteraemia in HIV-1 seropositive adults admitted to hospital in Nairobi, Kenya. Lancet **336:**545–549.
17. **Hacker, J., L. Bender, M. Ott, J. Wingender, B. Lund, R. Marre, and W. Goebel.** 1990. Deletions of chromosomal regions coding for fimbriae and hemolysins occur *in vitro* and *in vivo* in various extraintestinal *Escherichia coli* isolates. Microb. Pathog. **8:**213–225.
18. **Hacker, J., S. Knapp, and W. Goebel.** 1983. Spontaneous deletions and flanking regions of the chromosomally inherited hemolysin determinant of an *Escherichia coli* O6 strain. J. Bacteriol. **154:**1145–1152.
19. **Hacker, J., M. Ott, G. Blum, R. Marre, J. Heesemann, H. Tschäpe, and W. Goebel.** 1992. Genetics of *Escherichia coli* uropathogenicity: analysis of the O6:K15:H31 isolate 536. Zentralbl. Bakteriol. **276:**165–175.
20. **Hacker, J., G. Schmidt, C. Hughes, S. Knapp, M. Marget, and W. Goebel.** 1985. Cloning and characterization of genes involved in production of mannose-resistant, neuraminidase-susceptible (X) fimbriae from a uropathogenic O6:K15:H31 *Escherichia coli* strain. Infect. Immun. **47:**434–440.
21. **Herzer, P. J., S. Inouye, M. Inouye, and T. S. Whittam.** 1990. Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of *Escherichia coli*. J. Bacteriol. **172:**6175–6181.
22. **High, N. J., B. A. Hales, K. Jann, and G. J. Boulnois.** 1988. A block of urovirulence genes encoding multiple fimbriae and hemolysin in *Escherichia coli* O4:K12:H−. Infect. Immun. **56:**513–517.
23. **Hull, R., S. Bieler, S. Falkow, and S. Hull.** 1986. Chromosomal map position of genes encoding P adhesins in uropathogenic *Escherichia coli*. Infect. Immun. **51:**693–695.
24. **Hull, R. A., R. E. Gill, P. Hsu, B. H. Minshew, and S. Falkow.** 1981. Construction and expression of recombinant plasmids encoding type 1 or D-mannose-resistant pili from a urinary tract infection *Escherichia coli* isolate. Infect. Immun. **33:**933–938.
25. **Hull, R. A., S. I. Hull, and S. Falkow.** 1984. Frequency of gene sequences necessary for pyelonephritis-associated pili expression among isolates of *Enterobacteriaceae* from human extraintestinal infections. Infect. Immun. **43:**1064–1067.
26. **Johnson, J. R.** 1991. Virulence factors in *Escherichia coli* urinary tract infection. Clin. Microbiol. Rev. **4:**80–128.
27. **Johnson, J. R., S. L. Moseley, P. L. Roberts, and W. E. Stamm.** 1988. Aerobactin and other virulence factor genes among strains of *Escherichia coli* causing urosepsis: association with patient characteristics. Infect. Immun. **56:**405–412.
28. **Johnson, J. R., I. Ørskov, F. Ørskov, P. Goullet, B. Picard, S. L. Moseley, P. L. Roberts, and W. E. Stamm.** 1994. O, K, and H antigens predict virulence factors, carboxylesterase B pattern, antimicrobial resistance, and host compromise among *Escherichia coli* strains causing urosepsis. J. Infect. Dis. **169:**119–126.
29. **Knapp, S., J. Hacker, T. Jarchau, and W. Goebel.** 1986. Large, unstable inserts in the chromosome affect virulence properties of uropathogenic *Escherichia coli* O6 strain 536. J. Bacteriol. **168:**22–30.
30. **Korhonen, T. K., M. V. Valtonen, J. Parkkinen, V. Väisänen-Rhen, J. Finne, F. Ørskov, I. Ørskov, S. B. Svenson, and P. H. Mäkelä.** 1985. Serotypes, hemolysin production, and receptor recognition of *Escherichia coli* strains associated with neonatal sepsis and meningitis. Infect. Immun. **48:**486–491.
31. **Koronakis, V., M. Cross, B. Senior, E. Koronakis, and C. Hughes.** 1987. The secreted hemolysins of *Proteus mirabilis*, *Proteus vulgaris*, and *Morganella morganii* are genetically related to each other and to the alpha-hemolysin of *Escherichia coli*. J. Bacteriol. **169:**1509–1515.

Downloaded from http://iai.asm.org/ on October 13, 2015 by University of Queensland Library

32. **Kreft, B., S. Bohnet, O. Carstensen, J. Hacker, and R. Marre.** 1993. Differential expression of interleukin-6, intracellular adhesion molecule 1, and major histocompatibility complex class II molecules in renal carcinoma cells stimulated with S fimbriae of uropathogenic *Escherichia coli*. Infect. Immun. **61:**3060–3063.
33. **Labigne-Roussel, A. F., D. Lark, G. Schoolnik, and S. Falkow.** 1984. Cloning and expression of an afimbrial adhesin (AFA-I) responsible for P blood group-independent, mannose-resistant hemagglutination from a pyelonephritic *Escherichia coli* strain. Infect. Immun. **46:**251–259.
34. **Leffler, H., and C. Svanborg-Edén.** 1981. Glycolipid receptors for uropathogenic *Escherichia coli* on human erythrocytes and uroepithelial cells. Infect. Immun. **34:**920–929.
35. **Low, D., V. David, D. Lark, G. Schoolnik, and S. Falkow.** 1984. Gene clusters governing the production of hemolysin and mannose-resistant hemagglutination are closely linked in *Escherichia coli* serotype O4 and O6 isolates from urinary tract infections. Infect. Immun. **43:**353–358.
36. **Maslow, J. N., S. Brecher, K. S. Adams, A. Durbin, S. Loring, and R. D. Arbeit.** 1993. Relationship between indole production and the differentiation of *Klebsiella* species: indole-positive and -negative isolates of *Klebsiella* determined to be clonal. J. Clin. Microbiol. **31:**2000–2003.
37. **Maslow, J. N., M. E. Mulligan, K. S. Adams, J. C. Justis, and R. D. Arbeit.** 1993. Bacterial adhesins and host factors in the development and outcome of *Escherichia coli* bacteremia. Clin. Infect. Dis. **17:**89–97.
38. **Maslow, J. N., M. E. Mulligan, and R. D. Arbeit.** 1994. Recurrent *Escherichia coli* bacteremia. J. Clin. Microbiol. **32:**710–714.
39. **Maslow, J. N., A. M. Slutsky, and R. D. Arbeit.** 1993. The application of pulsed field gel electrophoresis to molecular epidemiology, p. 563–572. *In* D. H. Persing, T. F. Smith, F. C. Tenover, and T. J. White (ed.), Diagnostic molecular microbiology: principles and applications. American Society for Microbiology, Washington, D.C.
40. **Mendiola, M. V., and F. de la Cruz.** 1989. Specificity of insertion of IS91, an insertion sequence present in α-haemolysin plasmids of *Escherichia coli*. Mol. Microbiol. **3:**979–984.
41. **Montgomerie, J. Z., G. M. Kalmanson, and L. B. Guze.** 1979. Enterobactin and virulence of *Escherichia coli* in pyelonephritis. J. Infect. Dis. **140:**1013.
42. **Nei, M.** 1987. Molecular evolutionary genetics. Columbia University Press, New York.
43. **Opal, S. M., A. S. Cross, P. Gemski, and L. W. Lyhte.** 1990. Aerobactin and α-hemolysin as virulence determinants in *Escherichia coli* isolated from human blood, urine, and stool. J. Infect. Dis. **161:**794–796.
44. **Ørskov, F., and I. Ørskov.** 1984. Serotyping of *Escherichia coli*. Methods Microbiol. **14:**43–112.
45. **Ott, M., L. Bender, G. Blum, M. Schmittroth, M. Achtman, H. Tschäpe, and J. Hacker.** 1991. Virulence patterns and long-range genetic mapping of extraintestinal *Escherichia coli* K1, K5, and K100 isolates: use of pulsed-field gel electrophoresis. Infect. Immun. **59:**2664–2672.
46. **Plos, K., S. I. Hull, R. A. Hull, B. R. Levin, I. Ørskov, F. Ørskov, and C. Svanborg-Edén.** 1989. Distribution of the P-associated-pilus (*pap*) region among *Escherichia coli* from natural sources: evidence for horizontal gene transfer. Infect. Immun. **57:**1604–1611.
47. **Rhen, M., P. Klemm, and T. K. Korhonen.** 1986. Identification of two new hemagglutinins of *Escherichia coli*, *N*-acetyl-D-glucosamine-specific fimbriae and a blood group M-specific agglutinin, by cloning the corresponding genes in *Escherichia coli* K-12. J. Bacteriol. **168:**1234–1242.
48. **Selander, R. K., D. A. Caugant, H. Ochman, J. M. Musser, M. N. Gilmour, and T. S. Whittam.** 1986. Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. Appl. Environ. Microbiol. **51:**873–884.
49. **Selander, R. K., D. A. Caugant, and T. S. Whittam.** 1987. Genetic structure and variation in natural populations of *Escherichia coli*, p. 1625–1648. *In* F. C. Neidhardt, K. L. Ingraham, B. Magasanik, K. B. Low, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
50. **Selander, R. K., T. K. Korhonen, V. Väisänen-Rhen, P. H. Williams, P. E. Pattison, and D. A. Caugant.** 1986. Genetic relationships and clonal structure of strains of *Escherichia coli* causing neonatal septicemia and meningitis. Infect. Immun. **52:**213–222.
51. **Siitonen, A.** 1992. *Escherichia coli* in fecal flora of healthy adults: serotypes, P and type 1C fimbriae, non-P mannose-resistant adhesins, and hemolytic activity. J. Infect. Dis. **166:**1058–1065.
52. **Sneath, P. H. A., and R. R. Sokal.** 1978. Numerical taxonomy. W. H. Freeman & Co., San Francisco.
53. **Strathdee, C. A., and R. Y. C. Lo.** 1987. Extensive homology between the leukotoxin of *Pasteurella haemolytica* A1 and the alpha-hemolysin of *Escherichia coli*. Infect. Immun. **55:**3233–3236.
54. **Svanborg-Edén, C., L.-M. Bjursten, R. Hull, S. Hull, K.-E. Magnusson, Z. Moldovano, and H. Leffler.** 1984. Influence of adhesins on the interaction of *Escherichia coli* with human phagocytes. Infect. Immun. **44:**672–680.
55. **Temmerman, M., C. F. Gilks, and H. C. G. Sanghvi.** 1993. Spontaneous and induced abortions at Kenyatta National Hospital, Nairobi, Kenya. Int. J. Gynaecol. Obstet. **41:**182–183.
56. **Tullus, K., A. Brauner, B. Fryklund, T. Munkhammar, W. Rabsch, R. Reissbrodt, and L. G. Burman.** 1992. Host factors versus virulence-associated bacterial characteristics in neonatal and infantile bacteraemia and meningitis caused by *Escherichia coli*. J. Med. Microbiol. **36:**203–208.
57. **Väisänen-Rhen, V., J. Elo, E. Väisänen, A. Siitonen, I. Ørskov, F. Ørskov, S. B. Svenson, P. H. Mäkelä, and T. K. Korhonen.** 1984. P-fimbriated clones among uropathogenic *Escherichia coli* strains. Infect. Immun. **43:**149–155.
58. **Valvano, M. A., R. P. Silver, and J. H. Crosa.** 1986. Occurrence of chromosome- or plasmid-mediated aerobactin iron transport systems and hemolysin production among clonal groups of human invasive strains of *Escherichia coli* K1. Infect. Immun. **52:**192–199.

58a. **Whittam, T. S.** Unpublished data.

59. **Whittam, T. S., M. L. Wolfe, I. K. Wachsmuth, F. Ørskov, I. Ørskov, and R. A. Wilson.** 1993. Clonal relationships among *Escherichia coli* strains that cause hemorrhagic colitis and infantile diarrhea. Infect. Immun. **61:**1619–1629.
60. **Williams, P. H., and N. H. Carbonetti.** 1986. Iron, siderophores, and the pursuit of virulence: independence of the aerobactin and enterochelin iron uptake systems in *Escherichia coli*. Infect. Immun. **51:**942–947.

Downloaded from http://iai.asm.org/ on October 13, 2015 by University of Queensland Library